

# Sensitivity of Contingent Value Surplus Estimates to Elicitation Approach: Further Evidence

J.M. Bowker, R.A. Souter, and J.R. Clemmons

**Abstract.**—Two elicitation approaches, dichotomous choice (DC) and payment card (PC), were commonly used in contingent valuation (CV) studies to measure Hicksian economic surplus were examined. Separate samples were used to estimate individual annual net economic surplus associated with lake recreation in the Whiskeytown-Shasta-Trinity National Recreation Area (WST) with each elicitation approach. Results indicate a large and statistically significant difference in estimates obtained from PC and DC. Results indicate that CV estimates are not invariant to elicitation technique and that use of a single CV elicitation approach in applied benefit-cost analysis should be tempered with caution.

**Keywords:** CV, dichotomous choice, payment card, logit, incongruence.

---

## INTRODUCTION

Contingent valuation (CV) is a popular and accepted method for valuing entities not commonly traded in markets (Mitchell and Carson 1989; U.S. Water Resources Council 1983). The technique utilizes survey methods to directly elicit values from individuals for the provision of a particular good or service through the use of a hypothetically structured market or similar property right transfer institution. CV presents a flexible and convenient alternative to other nonmarket valuation techniques such as travel cost (Smith and others 1986), hedonic pricing (Brookshire and others 1982), and defense expenditures (Abdalla 1990), which are indirect methods reliant on the existence of a complementary or substitute relationship between the nonmarket good of interest and some other market good (Randall 1987).

CV has been applied to a wide range of nonmarket valuation problems. Examples include studies to estimate values for wilderness use and preservation (Gilbert and others 1992; Walsh and others 1984), water based recreation (Cordell and Bergstrom 1993; Sellar and others 1985), wildlife viewing and harvesting (Bishop and Heberlein 1979; Duffield and Patterson 1991; Peterson and others 1992), and public aerobic classes (McCarville 1991).

Criticisms of CV generally focus on the many biases which can result when the method is applied (McKillop 1992). (Mitchell and Carson 1989) provide a complete typology of these biases. Additional and perhaps more indictable problems with CV relate to individual valuation processes. For example, in an experiment valuing wildlife preservation, (Stevens and others 1991) found that a large proportion of respondents simply refused to attach a money metric to species preservation. (Kahneman and Knetsch 1992) argue that values obtained in CV studies often are potentially inflated due to two phenomenon: (1) "embedding" and/or (2) "purchase of moral satisfaction." In

---

The first two authors are a Social Scientist and a Mathematical Statistician at the Southeastern Forest Experiment Station of the USDA Forest Service, 320 Green St., Athens, GA 30605. The third author is a Computer Programmer at the University of Florida, Gainesville, FL.

addition, Samples and others (1986) have shown that sequencing of questions in CV can lead to disparate value measures and that in the case of wildlife, values for individual species, when summed exceed reported aggregate values.

Biases and other problems must be weighed against the alternative of using market information with its inherent distributional biases, or worse, assuming a zero price for a nonmarket good is efficient and legitimate. As Smith (1992) contends, "true economists can never know the 'true' values people place on any commodity—marketed or nonmarketed." Additionally, few would argue preferences and values are not somewhat dynamic.

Attempts at validation of the CV technique generally consist of convergent validity experiments. Empirical studies have been limited to the issues of: (1) temporal stability (Loomis 1990), (2) stability across elicitation procedures (Boyle and Bishop 1988), (3) comparing CV values to those obtained via indirect methods (Richards and others 1990; Sellar and others 1985; Smith and others 1986), and (4) comparing CV values to those obtained through simulated markets involving cash transactions (Bishop and Heberlein 1979; Dickie and others 1987; Kahneman and Knetsch 1992; Kealy and others 1990). While limited to a collection of case studies, the results generally suggest that when compared to indirect or simulated market values for similar goods, values estimated via CV are very reasonable, supporting convergent validity.

In spite of favorable published findings (Boyle and Bishop 1988; Loomis 1990), basic questions remain about the appropriateness of elicitation approaches employed within the CV method. Most studies select one method of elicitation, either open-ended (OE), payment card (PC), or dichotomous choice (DC). A benefit of using only one elicitation method is enhanced statistical efficiency for a given sample size, which is desirable given limited budgets and the expense of collecting survey data. A potential danger of using only one elicitation method is that elicited values may be incongruent across procedures and a level of precision may be represented which is artificially high. This occurrence may be problematic because there are no compelling reasons, theoretical or empirical, to select any one elicitation procedure as the most valid or accurate.

## DATA AND METHODS

Data used for the analysis were obtained via surveys conducted during the recreation seasons of 1991 and 1992 at Whiskeytown-Shasta-Trinity National Recreation Areas (WST). Dichotomous choice data were obtained in 1991, while the lakes were in severe drought condition, using a random sample of on-site visitors. In addition to a number of other questions, respondents were asked in a face-to-face interview to answer either yes or no to the following question aimed at eliciting annual household economic surplus obtained through using WST:

Think for a moment about the total amount of money spent during the last 12 months to visit WST. The "amount of money" is the total amount you spent on all trips to WST in the last 12 months including this trip. If the total amount you had to spend on all trips to WST last year had been \$ D higher, for example because of higher price of gasoline or other expenses, would you still have come?

The \$D amount ranged from \$1 to \$250 in the following increments: \$ 1, 2, 5, 7.5, 10, 15, 10, 15, 25, 45, 65, 85, 110, 150, 250. A follow-up question was asked to identify protest bidders. A total of 167 surveys were completed of which 3 were identified as protest bids and deleted. An additional 14 surveys were unusable because of failure to complete the questionnaire (13 of which were for the income question). The usable sample was 150 observations.

Because the DC approach is based on YES/NO responses, nonlinear binary response modeling (logit or probit) is used to estimate individual economic surplus (Hanemann 1984). The process is two-stage. First, the inverse

of the distribution function for willingness to pay (WTP), based on the probability of YES response, is estimated. The logit model is specified as,

$$P(\text{YES}) = 1 / (1 + \exp - (B_0 + B_1\$D + B_2\text{INC} + B_3\text{QUAL} + u)) \quad (1)$$

where, P(YES) is the probability of a YES response to an amount of \$D dollars in additional annual expenses to continue to use the site, INC is household income, QUAL is an index of site quality and u represents i.i.d. mean zero random perturbation. Next, conditional means and medians for WTP can be obtained via numerical techniques (Bowker and Stoll 1988; Hanemann 1984). Confidence intervals for the WTP means and medians may be estimated using (1) differential, (2) bootstrap, or (3) Monte Carlo techniques (Duffield and Patterson 1991).

Payment card data were obtained via a mail back survey distributed on-site in 1992. Site conditions such as weather and water levels during both years were virtually identical. In the survey, respondents were shown pictures of water levels corresponding to the annual drought conditions and asked to:

Consider the water levels as shown and described on page X. Consider also your estimated annual total household expenses to use WST during a drought year, for example gasoline, travel expenses, boat or cabin rentals and all other expenses. Please circle the number closest to the maximum you would pay in additional annual expenses in a drought year to continue using WST. \$ 0 2 5 10 . . . . 1,000.

In all, 33 amounts, ranging from \$0 to \$1,000 were listed. As with the DC question, a follow-up question was asked to identify protest bidders. Observations deemed as protest bids, less than 5 percent of the sample, were deleted. A total of 111 usable observations remained.

With the PC, WTP is obtained by directly asking individuals to indicate on a payment card their WTP in the form of additional annual expenses incurred to continue to use WST. In this case, WTP (or a transformation thereof) is modeled directly using ordinary least squares (OLS) (Cameron and Huppert 1989) as a linear function of socioeconomic and site variables, where, INC, QUAL, and u are as described above.

$$\text{WTP} = B_0 + B_1\text{INC} + B_2\text{QUAL} + u \quad (2)$$

From this specification, conditional mean and median WTP measures can be directly obtained. Confidence intervals for the mean may be obtained by conventional procedures using either the sample or regression standard error. Conover (1980) provides a nonparametric procedure to estimate a confidence interval for the median.

## RESULTS

Data for the payment card method indicate a sample mean WTP of \$15.67, and sample standard error of 22.76. In this case, two empirical specifications are used; the first with WTP as the dependent variable, and the second using the natural logarithm of WTP as the dependent variable, Table 1. The covariates for both specifications are income and a Likert-type quality index. Signs on the explanatory variable coefficients for both specifications are as expected; however, only the coefficient on the quality index variable is significant. These findings are consistent with other cross-sectional studies in that income often appears to be an insignificant factor in explaining WTP. Also consistent with a number of other cross-sectional studies is the low  $R^2$  indicating relatively poor predictive power.

Table 1—Payment card, ordinary least squares regression estimates

| Dependent variable | CONST<br>(Constant) | INC<br>(Income) | QUAL<br>(Site quality) | R <sup>2</sup> | F     | N   |
|--------------------|---------------------|-----------------|------------------------|----------------|-------|-----|
| WTP <sup>1</sup>   | -4.73               | 1.33            | 4.76                   | .134           | 8.34  | 111 |
| t's                | (-.79)              | (1.50)          | (3.60)                 |                |       |     |
| LNWTP <sup>2</sup> | -.273               | 0.019           | 0.563                  | .200           | 13.80 | 111 |
| t's                | (-.56)              | (0.26)          | (5.20)                 |                |       |     |

<sup>1</sup> Willingness to pay<sup>2</sup> Natural log of willingness to pay

Incorporating the regression standard error from the WTP model, the 95 percent confidence interval for mean WTP at the mean of the regressors is [-27.07, 58.41]. A well known property of the semi-logarithmic specification is that the antilog of the mean LNWTP provides an estimate of the median WTP. Using the regression standard error for the LNWTP model, the 95 percent confidence interval for median WTP at the mean of the regressors is [0, 134].

The estimated MLE logit regression equation from the dichotomous choice approach is presented in table 2. While the model predicts YES/NO responses correctly on 94 percent of the observations, the McFadden R<sup>2</sup> (MDR<sup>2</sup>) is relatively low and the chi-squared statistic for the nonintercept parameters in the model is significant at the 10 percent level. This indicates a relatively poor fit which obtains from the fact that over 80 percent of the respondents said YES to the \$D amount on their question and the proportions changed little across the range of WTP amounts. The sign on the offer amount variable, \$D, is theoretically correct but statistically insignificant while income is insignificant and quality is significant at the 5 percent level.

Table 2—Dichotomous choice logit MLE regression estimates

| DEP<br>(Dependent variable) | CONST<br>(Constant) | \$D<br>(offer amount) | INC<br>(Income) | QUAL<br>(Site Quality) | MDR <sup>2</sup> | CHI <sup>2</sup> | N   |
|-----------------------------|---------------------|-----------------------|-----------------|------------------------|------------------|------------------|-----|
| P(YES)                      | 10.21               | 0.0047                | 0.163           | -1.26                  | 0.10             | 6.38             | 150 |
| t's                         | (2.5)               | (-1.1)                | (.849)          |                        |                  |                  |     |

Point estimates of the truncated mean WTP, obtained via numerical integration over ranges of offer amounts of 250 and 1000 were \$239.20 and 728.97 respectively. Alternatively, following Hanemann (1984), the parameter estimates of the above logit function can be combined to yield an estimated mean and median WTP of \$790.23. A 95- percent confidence interval for the mean WTP obtained via the Krinsky and Robb Monte Carlo simulation approach is [\$315.88, 14151.95].

Our findings show the DC WTP estimates to be substantially higher than those obtained via PC. Indeed, 95 percent confidence intervals for the mean WTP out of the two elicitation approaches do not intersect. This would appear to be sufficient evidence to reject the hypothesis that  $WTP_{PC} = WTP_{DC}$ . Our results run contrary to the findings of Loomis (1990) and Boyle and Bishop (1988), by showing that CV can be quite sensitive to elicitation vehicle selection. While our study may be an anomaly, we see no reason why our results are any more an artifact of a single strange case than the findings of either Loomis (1990) or Boyle and Bishop (1988).

## IMPLICATIONS

These results suggest that researchers using CV to estimate recreation values and benefits should consider the danger of using only one elicitation procedure versus the loss of statistical efficiency associated with splitting the sample over two or more elicitation procedures. Given that procedure selection criteria are basically ad hoc, we would argue that selection of a single approach conveys a level of precision and accuracy which is artificially high. Indeed, the DC mean WTP is 15 times higher than that for the PC.

Our findings must be tempered with a number of caveats. We recognize that a number of experimental control factors are likely to have contributed to the incongruence of our WTP estimates. For example, (1) the data generating processes were based on different questions in different years, (2) interviewing was face-to-face with DC and via mailback with PC, (3) visuals were used only with PC. However, we believe our empirical findings indicate something more fundamental; an inherent instability of WTP estimates to elicitation approach variations. An important implication of this finding is that when CV is used in public policy decision processes such as applied benefit-cost analysis, vastly different economic surplus estimates and consequent conclusions may emerge depending on the elicitation vehicle. If the results are to be aggregated across user populations (Loomis 1987) and used in the policy arena, they should be viewed and interpreted cautiously. Clearly, peoples' preferences are dynamic and measuring those preferences in a money metric is subject to error. However, incongruence of the magnitude reported in this study is hard to overlook. Such a finding appears to corroborate much of the criticism that the CV technique is currently receiving.

## REFERENCES

- Abdalla, C.W. 1990. Measuring economic losses from groundwater contamination: an investigation of household avoidance costs. *Water Resources Bulletin*. 26(3):451-463.
- Bishop, R.C.; Heberlein, T.A. 1979. Measuring values of extra market goods: are indirect measures biased? *American Journal of Agricultural Economics*. 61(5):926-930.
- Bowker, J.M.; Stoll, J.R. 1988. Use of dichotomous choice nonmarket methods to value the whooping crane resource. *American Journal of Agricultural Economics*. 70(2):372-381.
- Boyle, K.J.; Bishop, R.C. 1988. Welfare measurements using contingent valuation: a comparison of techniques. *American Journal of Agricultural Economics*. 70(1):20-28.
- Brookshire, D.S.; Thayer, M.A.; Shultze, W.D.; D'Arge, R.C. 1982. Valuing public goods: a comparison of survey and hedonic approaches. *The American Economic Review*. 72(1):165-176.
- Cameron, T.A.; Huppert, D.D. 1989. OLS versus ML estimation of nonmarket resource values with payment card interval data. *Journal of Environmental Economics and Management*. 17(3):230-246.
- Conover, W.J. 1980. *Practical nonparametric statistics*. 2nd. ed. New York: John Wiley and Sons.
- Cordell, H.K.; Bergstrom, J. 1993. Comparison of recreation use values among alternative reservoir water-level management scenarios. *Water Resources Research*. 29(2):247-258.
- Dickie, M.; Fisher, A.; Gerking, S. 1987. Market transactions and hypothetical demand data: a comparative study. *Journal of the American Statistical Association*. 82(397):69-75.
- Duffield, J.W.; Patterson, D.A. 1991. Inference and optimal design for a welfare measure in dichotomous choice contingent valuation. *Land Economics*. 67(2):225-239..

- Gilbert, A.; Glass, R.; More, T. 1992. Valuation of eastern wilderness: extra market measures of public support. In: Payne, C., Bowker, J.M.; and Reed, P., eds. *The economic value of wilderness: proceedings of the conference*. Gen. Tech. Rep. SE-78. Asheville, NC: Southeastern Forest Experiment Station: 57-70.
- Hanemann, M. 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*. 66(3):332-341.
- Kahneman, D.; Knetsch, J.L. 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management*. 22:57-70.
- Kealy, M.J.; Montgomery, M.; Dovidio, J.F. 1990. Reliability and predictive validity of contingent values: does the nature of the good matter? *Journal of Environmental Economics and Management*. 19(1):244-263.
- Loomis, J.B. 1987. Expanding contingent value sample estimates to aggregate benefit estimates: current practices and proposed solutions. *Land Economics*. 63(4):396-402.
- Loomis, J.B. 1990. Comparative reliability of the dichotomous choice and open-ended contingent valuation techniques. *Journal of Environmental Economics and Management*. 18(1):75-85.
- McCarville, R.E. 1991. An empirical investigation of the influence of cost information on willingness to pay for public aerobics classes. *Leisure Sciences*. 13(1):85-96.
- McKillop, W. 1992. Use of contingent valuation in northern spotted owl studies: a critique. *Journal of Forestry*. 9(80):36-37.
- Mitchell, R.C.; Carson, R.T. 1989. *Using surveys to value public goods: the contingent valuation method*. Washington, DC: Resources for the Future. 462 p.
- Peterson, G.L.; Sorg, C.S.; McCollum, D.W.; Thomas, M.H. 1992. *Valuing wildlife in Alaska*. Boulder, CO: Westview Press, Inc. 357 p.
- Randall, A. 1987. *Resource economics—an economic approach to natural resource and environmental policy*. 2nd. ed. New York: John Wiley and Sons.
- Richards, M.T.; King, D.A.; Daniel, T.A.; Brown, T. 1990. The lack of an expected relationship between travel cost and contingent value estimates of forest recreation. *Leisure Sciences*. 12:303-319.
- Samples, K.C.; Dixon, J.A.; Gowan, M.M. 1986. Information disclosure and endangered species evaluation. *Land Economics*. 62(August):306-312.
- Sellar, C.; Stoll, J.R.; Chavas, J.P. 1985. Validation of empirical measures of welfare change: a comparison of nonmarket techniques. *Land Economics*. 61(2):156-175.
- Smith, V.K. 1992. Arbitrary values, good causes, and premature verdicts: comment. *Journal of Environmental Economics and Management*. 22:71-89.
- Smith, V.K.; Desvousges, W.H.; Fisher, A. 1986. A comparison of direct and indirect methods for estimating environmental benefits. *American Journal of Agricultural Economics*. 68:280-290.
- Stevens, T.H.; Glass, R.; More, T.; Echeverria, J. 1991. Wildlife recovery: is benefit-cost analysis appropriate? *Journal of Environmental Management*. 33(3):327-334.

U.S. Water Resources Council. 1983. Economic and environmental principles and guidelines for water and related land resources implementation studies. Washington, DC: U.S. Government Printing Office.

Walsh, R.G.; Loomis, J.; Gillman, R.S. 1984. Valuing option, existence and bequest demands for wilderness. *Land Economics*. 60(1):14-29.